AD-785 367

APPLICATION OF RESPONSE SURFACE METHODOLOGY TO SHAPE DISCRIMINATION

Lawrence A. Scanlan

Illinois University

Prepared for:

Air Force Office of Scientific Research

February 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

UNCLASSIFIED
Security Classification

AD 785 367

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Aviation Research Laboratory Institute of Aviation University of Illinois, Savoy, Illinois	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP
3. REPORT TITLE APPLICATION OF RESPONSE SURFACE METHODOLOGY TO SHAPE DISCRIMINATION	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific Interim	
5. AUTHOR(S) (First name, middle initial, last name) Lawrence A. Scanlan	
6. REPORT DATE February 1974	7a. TOTAL NO. OF PAGES 37 7b. NO. OF REFS 41
8a. CONTRACT OR GRANT NO. F44620-70-C-0105 b. PROJECT NO. 9778 c. 61102F d. 681313	9a. ORIGINATOR'S REPORT NUMBER(S) ARL-74-2/AFOSR-74-1 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR - TR - 74 - 1445
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.	
11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research 1400 Wilson Boulevard (NL) Arlington, Virginia 22209

13. ABSTRACT

The effects of systematically manipulating five physical features of random, closed figures were determined. Subjects were asked to select one of four figures presented as being least like the others. The five features investigated were: number of edges, area, radial length variance, edge length variance, and orientation of the major axis. A five-factor, five-level, central-composite response surface design was employed. The resulting regression equation indicated that four features were reliable in their prediction of performance. Additionally several interaction effects were found. The applicability of these data for computer simulation is discussed.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Response Surface Methodology						
Form Perception						
Central Composite Designs						
Computer Simulation						

AVIATION RESEARCH LABORATORY

INSTITUTE OF AVIATION
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
WILLARD AIRPORT
SAVOY, ILLINOIS
61874

TECHNICAL REPORT

ARL-74-2
AFOSR-74-1
FEBRUARY 1974

APPLICATION OF RESPONSE SURFACE METHODOLOGY TO SHAPE DISCRIMINATION

LAWRENCE A. SCANLAN

PREPARED FOR

LIFE SCIENCES PROGRAM
OFFICE OF SCIENTIFIC RESEARCH
AIR FORCE SYSTEMS COMMAND

F44620-70-C-0105

FOREWORD

This research was sponsored by the Life Sciences Program, Air Force Office of Scientific Research, Contract Number F44620-70-C-0105. Dr. Charles E. Hutchinson was the scientific monitor of the contract. The guidance in the formulation of the problem provided by Dr. Bruce McCormick, Head, Information Engineering Department, University of Illinois, Chicago Circle was greatly appreciated.

1C

ABSTRACT

The effects of systematically manipulating five physical features of random, closed figures were determined. Subjects were asked to select one of four figures presented as being least like the others. The five features investigated were: number of edges, area, radial length variance, edge length variance, and orientation of the major axis. A five-factor, five-level, central-composite response surface design was employed. The resulting regression equation indicated that four features were reliable in their prediction of performance. Additionally several interaction effects were found. The applicability of these data for computer simulation is discussed.

TABLE OF CONTENTS

INTRODUCTION

The question of how man perceives form has held a prominent position in psychological thinking for a great many years. Beginning with William James (1890) countless psychologists have philosophized, theorized, and experimented on man's ability to perceive form. Zusne's (1970) review of the psychological literature on form perception lists 2583 references, and the entire Gestalt school of psychology is based on principles derived from form perception. In addition to the work done by psychologists, a great many computer scientists and engineers have worked on similar problems encountered in the attempt to perform pattern recognition by machine. It is discouraging to note that the level of our understanding of man's perceptual process is apparently inversely related to the amount of work that has been done. Selfridge and Neisser (1960) have noted that the inability to simulate perception has severely limited progress towards computer simulation of cognitive processes. As they note "... until programs to perceive patterns can be developed, achievements in mechanical problem solving will remain isolated technical triumphs [p. 60]."

As Zusne's (1970) review indicates the difficulty is not a lack of psychological experimentation per se but rather a lack of psychological models amenable to computer simulation. Theoretical formulations, such as the Gestalt principles of "good form" (law of Praegnanz), "good continuation," and "closure" (Koehler, 1929; Koffka, 1935), while clearly demonstrable, do not lend themselves to easy implementation in a computer simulation.

Similarly data obtained using multidimensional scaling techniques (Kunnapos, Malhammer, and Svenson, 1964; Nunnally, 1967; Torgerson, 1958), which attempt to map the physical dimensions of figures (features) into a psychological space, have failed to provide the information required for automatic pattern recognition. As an example of the method, Stenson (1968) computed a number of physical features for each of 20 figures then obtained subject ratings of judged similarity for pairs of

figures. These data were analyzed using multidimensional scaling methods and identified four factors (complexity, curvature, curvature dispersion, and straight-length dispersion) which accounted for the majority of the variance. A number of problems arise in this type of experimentation. First, the sample size of figures is generally quite small yielding data which could potentially be unreliable. Second, it is necessary to know the characteristics of the parent population of figures from which the sample of figures is drawn (Brown and Michels, 1966; Brown and Owen, 1967). Third, there is no agreement as to which scaling method is most appropriate (Attneave, 1950; Hake, 1966). Additionally, the procedure does not yield appropriate information about the relative importance of the resulting dimensions and, more importantly, no information about the interaction of two or more features is provided. Indeed, it is entirely possible that an important physical feature might not appear to be important because of its interaction with other features. A multiple regression approach, where subject performance is described by an equation containing a series of weighted terms representing individual physical features and combinations of features, would provide the type of data needed for cognitive simulation. The research of Maruyama (1971) is an example of one technique of obtaining what is essentially a linear regression equation and because it served as the impetus for the study to be reported here will be reviewed in some detail.

Maruyama was interested in obtaining a computer program that could choose which of several presented shapes was least like the others (odd shape detection). To accomplish this, he computed 22 features for each of the presented shapes. For each feature, he determined the shapes that had the largest and smallest feature value. The odd figure was then determined by finding the shape that was extreme on the largest number of features. Initial comparisons of computer performance and human subject performance on a set of 20 test trials indicated a relatively poor correspondence. To make the machine choice more like average subject choice required that certain features be more heavily weighted

than others. This results in a procedure that amounts to the evaluation of a linear regression equation of the form:

$$y_k = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{22} x_{22}$$

where x_1 is the deviation of feature 1 for figure k from the average of feature 1 for all of the other figures being compared. The shape with the largest y would then be chosen as the odd shape. Maruyama tried a number of weighting factors and chose the one producing the closest correspondence with his subject data.

Although this basic approach of using a regression equation appears to hold promise, a number of problems with Maruyama's procedure can be identified. The first problem involves the use of a simple multiple regression formula as opposed to a polynomial multiple regression equation which would include higher order terms representing interactions among features. The second problem lies in the method used to determine weights for the various features. Several potential forms of weighting functions were defined and then linear programming techniques were used to obtain the best fitting coefficients using subject data as the criterion. The weighting function yielding the best correspondence with the subject responses was then kept. An obvious deficiency is that only a small number of all possible weighting functions can be tested. This approach also requires a sample of test figures that is large enought to represent adequately the population of figures to which the final model is to apply. Obviously, with only 20 trials the sampling error will be too large for the final model to be reliable.

Potentially a much better approach would be to obtain subject data in such a manner that a polynomial multiple regression equation could be obtained and an analysis of variance performed on the beta weights to determine which of them reliably contribute to the prediction of subject performance. The use of polynomial regression allows the inclusion of the potentially important interactive effects, and the ability to assess the reliability of each term allows the resulting model to retain only relevant effects.

Such a procedure was implemented in the present experiment. Prior to experimentation, however, steps had to be taken to reduce the impossibly large amount of data required to fit a complete polynomial for 22 features. Two steps were taken to affect this reduction. First, the number of features was reduced to five by carefully selecting features likely to be relevant. Second, a response surface methodology experimental design was employed. The next two sections detail these steps and are followed by a description of the procedures used to obtain test figures.

SELECTION OF FEATURES

A number of criteria were involved in the choice of which features would be manipulated in the current experiment. Because of the characteristics of the experimental design to be used, it was necessary that the features be orthogonal. (That is, all possible levels of one feature could occur with any combination of levels of the other features). The total number of features needed to be relatively small to keep the number of experimental conditions reasonable. Previous research and the features used there, needed to be considered wherever possible.

Because the task to be employed in the current study was to be the same as that used by Maruyama (1971), the first step in the selection of features involved a careful look at Maruyama's data. As indicated previously, Maruyama employed 22 features in his program, however, many of these were highly correlated with one another. For example, perimeter divided by area, perimeter squared divided by area, and perimeter divided by the square root of the area are all very closely related to one another. As an orthogonal set of features was required for this study, additional information was necessary to determine which orthogonal subset of Maruyama's features would be appropriate.

This additional information was obtained by asking a group of six subjects to examine Maruyama's figures and choose the odd shape and the reason for their choice. An examination of the reasons given for a particular choice of odd shape revealed size, number of edges, and non-regularity of shape as the most commonly given reasons. Next, for those sets of figures where at least five of the subjects selected the same figure as odd, those features with the most extreme values were identified. These are the features that would most influence the computer's choice of odd shape. Of the 20 sets of figures shown to the subjects, nine met the above criterion. For these nine sets the most influential features were the number of verticies or edges, the deviation in edge length, the radius of a circle of equal area and the average deviation from this circle, the perimeter

(P) divided by the area (A), P^2/A, $P/\sqrt{A}$, and area. In terms of independent features the above set would reduce to the number of edges, the deviation in edge length, the area, and some measures of jaggedness.

The choice of an odd shape is basically a form discrimination problem, and, therefore, the form perception literature employing discrimination tasks should be applicable. However, because the interactive effects of several features varying simultaneously have not been systematically investigated, the literature is not as helpful as would be desired. The number of edges in random polygons (Brown, Hitchcock, and Michels, 1962; Coules and Lekarczyk, 1963; Crook, 1957) and dispersion, symmetry, and elongation (Andrews and Brown, 1967; Boynton, Elworth, Monty, Onley, and Klingberg, 1961; Boynton, Elworth, Onley, and Klingberg, 1960; Brown and LoSasso, 1967; Monty and Boynton, 1962; Zusne, 1970) have been shown to affect form discrimination. Although little work has been done on area as a feature, largely because a square is a square whether it is a large square or a small square, several studies have demonstrated that the estimation of area is affected by shape (Anastasi, 1936; Bolton, 1897; Warren and Pinneau, 1955). The direction and magnitude of the effect is not agreed upon, however. Although these studies have employed simple geometric figures rather than random shapes, the possibility of area interacting with other features appears very likely.

Data from scaling experiments suffer from a number of problems, some of which have been mentioned previously. Because of the diversity of tasks, figures, and scaling methods used, it is difficult to compare one study with another (see Zusne, 1970, 277-288). It is probably safe to say, however, that measures of jaggedness, complexity, compactness, and symmetry are likely to be important. A common measure of jaggedness is P^2/A, although it can be demonstrated that the variation in radial lengths from the centroid to the verticies is very nearly the same measure. Complexity is closely related to the number of edges although the variance of the edge length and the variance of the radial length might also influence judged complexity. An interaction of radial length variance and edge length variance would likely affect judgment of compactness.

The angular orientation of a figure in some cases will have no effect on the perception of a figure, as would be the case if a square is rotated a multiple of 90 degrees about its centroid. On the other hand, some changes in orientation create dramatic changes in perception, as would be the case if a square were rotated 45 degrees so that it is a diamond. Perhaps because of this equivocalness, rotation has received little experimental study. Because of this lack of data, the orientation of the longest axis from a vertex through the centroid was considered to be an important feature for study.

Based on the above considerations, a likely set of features to be included in the present study would be the number of edges, the area, the variance of the radial lengths, the variance of the edge lengths, and the angular orientation of the major axis of the figure.

EXPERIMENTAL DESIGN

The reduction in the number of features from 22 to 5 obviously reduces the number of data collection combinations dramatically. However, even with five features the number of required data points for a factorial experiment can be quite large. If each of the five features were tested at five different levels, the number of treatment combinations would be 3125 for each replication. Whereas it might be possible to test each of these combinations, alternative experimental designs that could reduce this number would be extremely attractive. One such alternative is response surface methodology (RSM) which was originally developed for use in the chemical industry (Box and Behnken, 1960; Box and Hunter, 1957; Box and Wilson, 1951). Because human subjects are considerably different in response variability than are chemical reactions, current RSM work has investigated the necessary modifications to the original RSM designs to allow their use in behavioral research (Clark and Williges, 1971; 1973; Simon, 1970; Williges and Simon, 1971).

As noted earlier the interaction of one variable with another can have considerable effect on subject response. However, when one begins examining the higher-order interactions it becomes apparent that the additional precision obtained by including the fourth- and fifth-order interactions in most cases does not warrant the additional data collection. RSM designs take advantage of this situation to reduce the data collection, while retaining the ability to assess the lower-order interactions. Additionally the analysis procedure (Clark, Williges, and Carmer, 1971) allows a determination of whether or not a higher-order regression equation (say one that contains third-order interactions) is necessary to describe the data. If higher-order terms are necessary, the additional data can be collected and a new regression equation can be obtained.

The particular design used in this study was a five-factor, five-level central-composite design consisting of 27 variable combinations. The variable level combinations are given in Mills and Williges (1973).

This particular design allows the determinction of a complete second-order polynomial regression equation of the form:

$$
\begin{aligned}
y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \\
& + \beta_6 x_1^2 + \beta_7 x_2^2 + \beta_8 x_3^2 + \beta_9 x_4^2 + \beta_{10} x_5^2 \\
& + \beta_{11} x_1 x_2 + \beta_{12} x_1 x_3 + \beta_{13} x_1 x_4 + \beta_{14} x_1 x_5 \\
& + \beta_{15} x_2 x_3 + \beta_{16} x_2 x_4 + \beta_{17} x_2 x_5 + \beta_{18} x_3 x_4 \\
& + \beta_{19} x_3 x_5 + \beta_{20} x_4 x_5 ,
\end{aligned}
$$

where

x_1 is the number of edges,

x_2 is the area,

x_3 is the variance of the radial lengths,

x_4 is the variance of the edge lengths, and

x_5 is the angular orientation of the major axis.

The x values can be either in coded form or in real-world values. For the design being used the coded values are -2, -1, 0, 1, and 2. The real-world values corresponding to these coded values will be discussed in the next section.

An analysis of variance performed on the above equation allows a determination of which of the 20 beta weights reliably predicts the subject data. After consideration of shrinkage discussed by Williges and North (1973) the entire design was replicated to obtain estimates of error variance.

A total of ten subjects served in the experiment. Each saw ten trials in each of the 27 treatment combinations for a total of 270 odd figure judgments. For each treatment condition the number of trials on which the subject selected the test figure was recorded allowing a determination of the probability of detection as the dependent variable.

To obtain the regression equation given above requires that the test figure differ from the standard comparison figures only on those features being manipulated. This means that of the four figures shown to a subject on any trial, three should be identical with respect to the five features under study. (This of course does not require that the standard figures "look" the same, and indeed they do not.) For the present study the standard figures were selected to have all five features at their center value. In terms of coded values all features were at the 0 level so that the regression equation reduces to $y = \beta_0$ when all four figures have the same standard features.

Specification of the experimental design determined the characteristics of the 270 test figures and the 810 standard figures. Generation of these figures required a precise definition of the features and the development of algorithms for computer solution.

STIMULUS GENERATION

A number of methods for the generation of random figures exist (Attneave and Arnoult, 1956), however, most of these are more appropriate for hand calculation than for computer generation. As the number of figures required for the present study was quite large, computer generation seemed justified. The figures used were random in the sense that the coordinates of the verticies were randomly chosen from a 256 x 256 matrix of points. This excluded a random radial and radial length approach as suggested by Thurmond (1966). Because one of the features to be calculated from these figures involved the radial distance from the centroid of the figure to each vertex, it was necessary to constrain the figures in such a way that all verticies were visible from the centroid. If such is not the case, some radials may intercept an edge before a vertex resulting in more than one possible value of radial length. Although an arbitrary choice could be made in such cases, instead the figure population was limited to angularly simple figures. (See Maruyama (1971) for a discussion of angularly simple figures and encoding processes).

Figure Algorithm

The following algorithm was employed to generate random figures:

1. Specify number of edges or vertices (NV)
2. Get a random x and y value for each of the NV verticies. (Two psudo-random number generators were used, one for x and one for y).
3. Calculate the center of mass for the vertex points. (This serves as a first approximation of the centroid).
4. Move the vertex points so that the center of mass is the 0, 0 point.
5. Calculate polar coordinate radial angles for each vertex.

6. Arrange the vertex points in descending order of their radial angles. (By so doing the i^{th} vertex can connect to the +1 (Modulo NV)th vertex with no crossing edges.)
7. Calculate the centroid of the figure and new radial angles.
8. Check to see that the radial angles are still in descending order. If not, the figure is not angularly simple and must be rejected.

Step 7 above involves calculation of the centroid which was obtained using a stepwise procedure. To find the x value of the centroid, the left-most vertex is found and used as a starting point. Moving to the right in steps of one, the area to the left of a vertical line is accumulated. When the area is half of the total area, the x value of the vertical line is the x coordinate of the centroid. A similar procedure is used for the y coordinate.

The incremental area is calculated by determining the intersections of all edges of the figure and the vertical line. Because there are checks on the line intersecting a vertex, the number of resulting intersections must be an even number. The table of intersections is arranged in descending order and the differences between the first and second and third and fourth, etc. intersections are added to the accumulated area.

The total area of the figure is determined by considering it as the sum of NV triangles with one common vertex at 0, 0. The general equation for the area of a triangle with verticies x_1y_1; x_2y_2; and x_3y_3 is:

$$A = 1/2 \left| x_1y_2 - x_2y_1 + x_2y_3 - x_3y_2 + x_3y_1 - x_1y_3 \right| .$$

If $x_3 = 0$ and $y_3 = 0$ as would be the case if one common vertex is at 0, 0 then the previous equation reduces to:

$$A = 1/2 \left| x_1y_2 - x_2y_1 \right| ,$$

and the area of a figure with NV edges would be:

$$A_{fig} = 1/2 \left| \left(\sum_{i=1}^{NV-1} (x_i y_{i+1} - x_{i+1} y_i)\right) + (x_{NV} y_1 - x_1 y_{NV}) \right| .$$

Feature Levels

For each of the five features, five levels could be accomodated by the experimental design. The choice was made to keep the figures relatively simple, therefore, the levels were chosen as 3, 4, 5, 6, and 7 edges. The major axis orientation variable levels were selected as 90 degrees right, 45 degrees right, vertical, 45 degrees left, and 90 degrees left. The area was manipulated in increments of 2000 units from 4000 to 12000.

The radial length variance and edge length variance features were calculated in a normalized form to make them independent of the area. The standard form of a variance is:

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{X})^2}{N}$$

where x_i is the individual data point, $\bar{X}$ is the mean, and N is the number of points. This particular equation is sensitive to the value of the mean, $\bar{X}$. If the mean is doubled, as would be the case if the size of the figure were doubled, then the variance is quadrupled. To compensate for this, the equation for variance was divided by $(\Sigma x_i)^2$ resulting in the following equation

$$\sigma^2_{norm} = \frac{NV^* \sum_{i=1}^{NV} x_i^2}{(\Sigma x_i)^2} - 1$$

where x_i is the i^{th} radial or edge length.

In the absence of any previous data to suggest what levels of these two variance measures to select, the choice was made to use the mean, the mean $\pm$ .5 standard deviation of the variance, and the mean ± 1.0 standard deviation. To determine the values of the mean and standard deviation of the distribution of the features, 1000 figures were generated for each of the five numbers of edges and the required estimates were obtained. Table 1 gives the results of these calculations. Table 2 summarizes the independent variable levels. Figures 1 to 5 show typical figures with each of the features manipulated over its five levels. In these figures the remaining four features are at their center value.

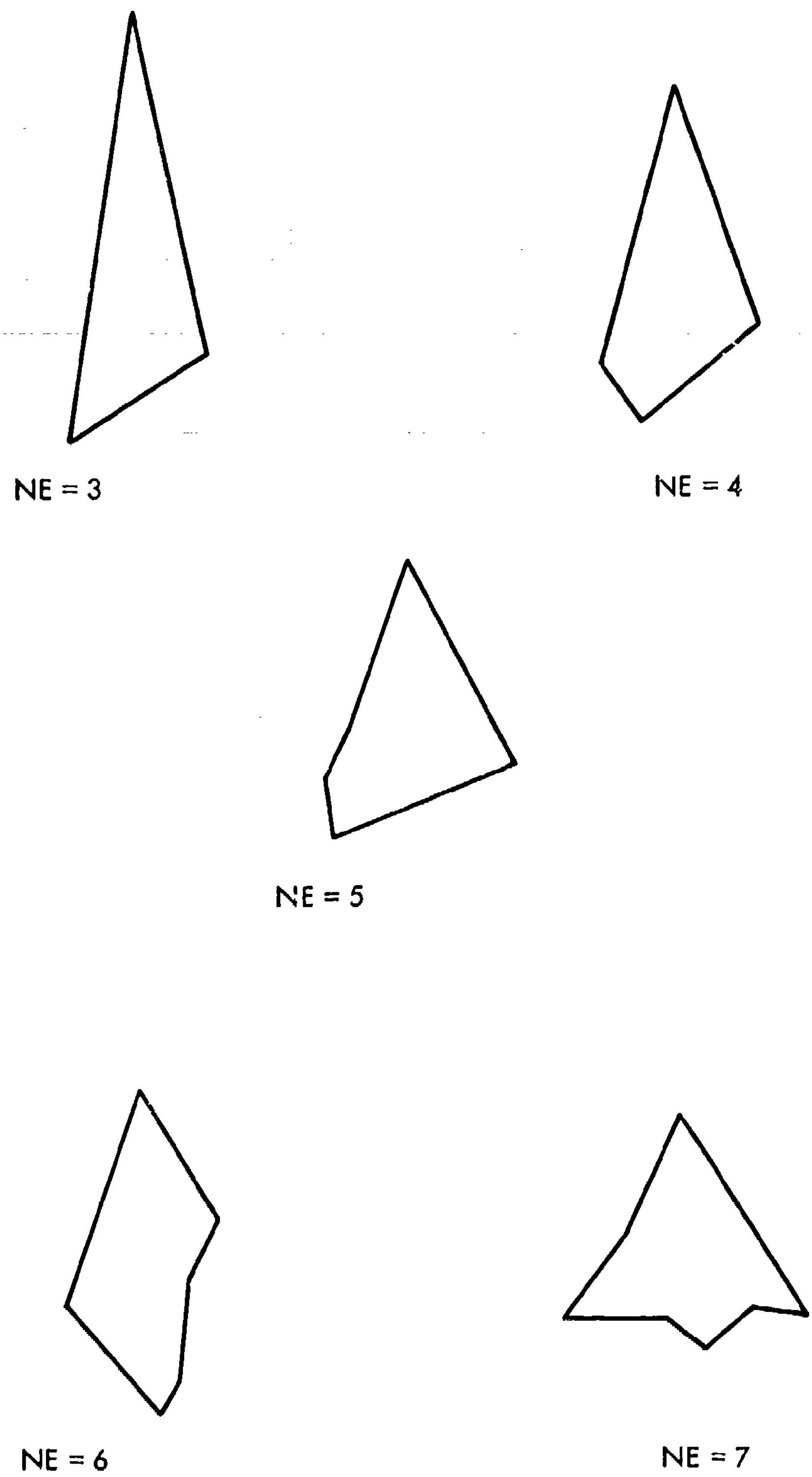


Figure 1. Typical test figures showing the effect of number of edges (NE).

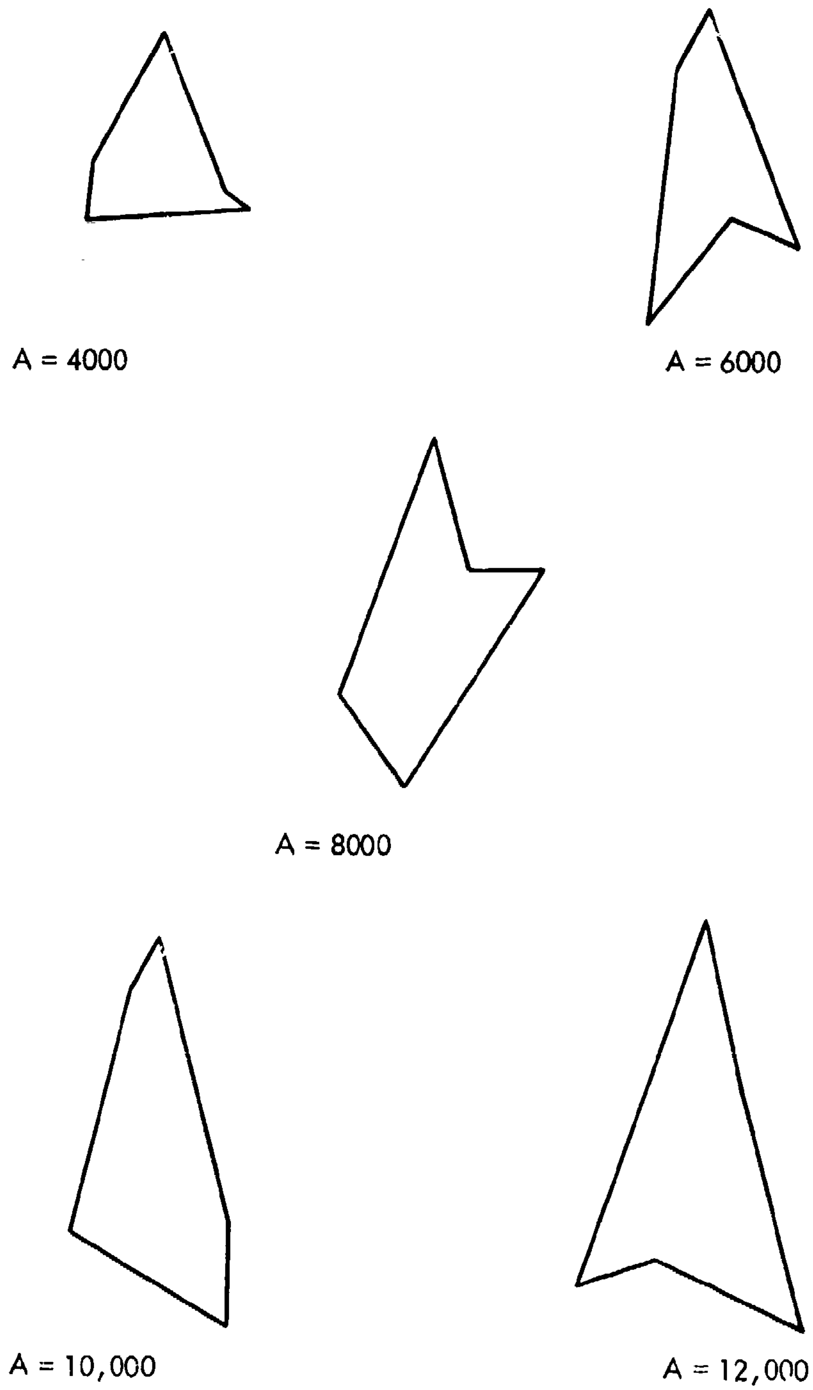


Figure 2. Typical test figures showing the effect of area (A).

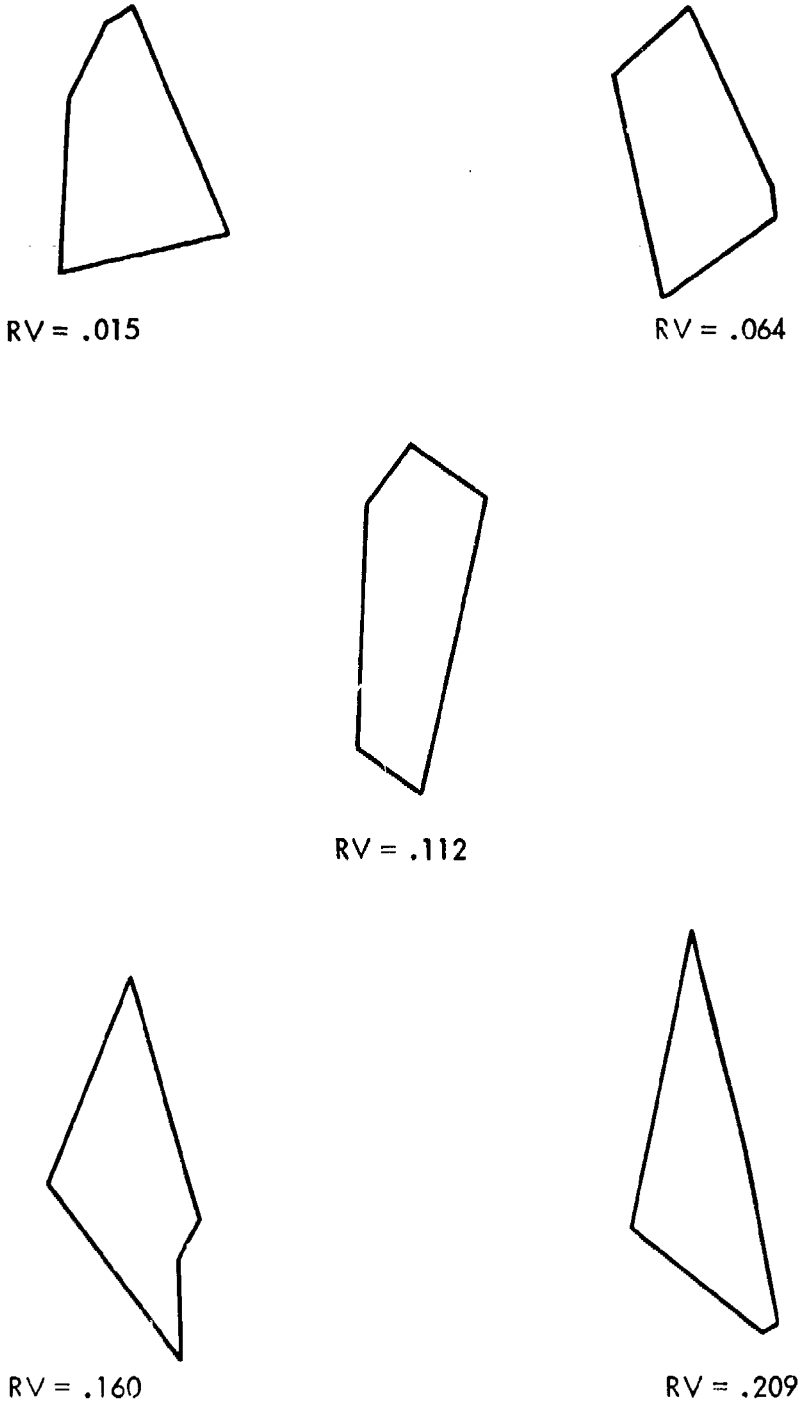
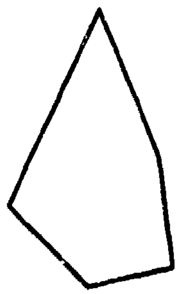
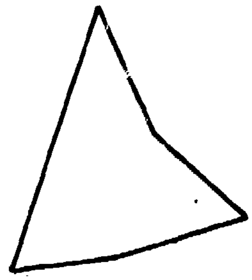


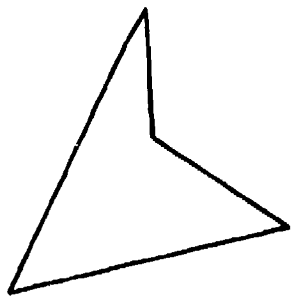
Figure 3. Typical test figures showing the effect of radial length variance (RV).



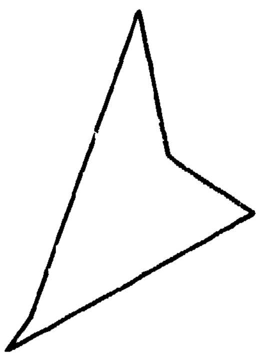
EV = .099



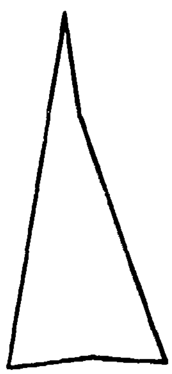
EV = .178



EV = .257



EV = .336



EV = .415

Figure 4. Typical test figures showing the effect of edge length variance (EV).

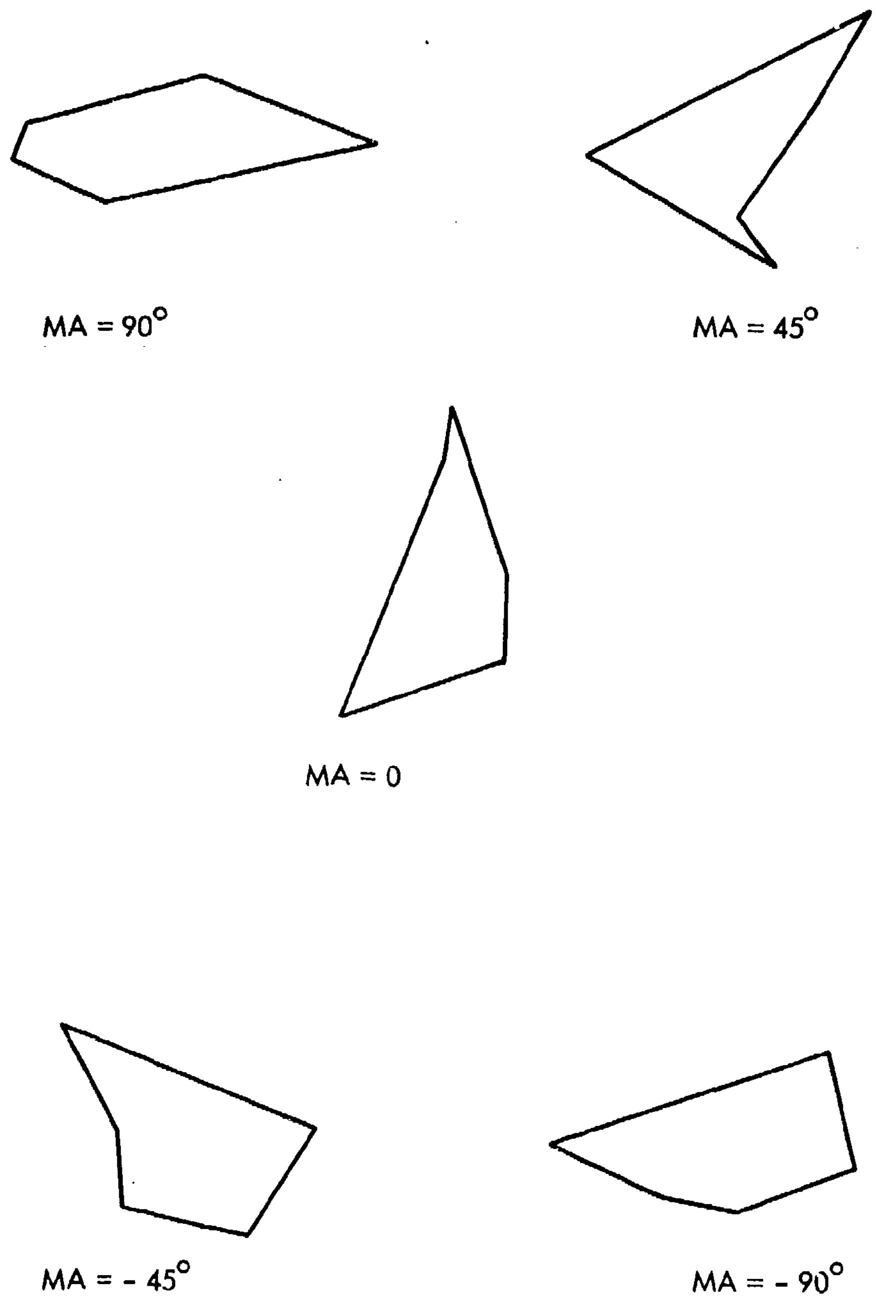


Figure 5. Typical test figures showing the effect of major axis orientation (MA).

TABLE 1

Mean and Standard Deviation of Radial Length Variance and Edge Length Variance Features as a Function of Number of Edges

Number of Edges	Radial Length Variance		Edge Length Variance	
	Mean	Std. Dev.	Mean	Std. Dev.
3	.157	.107	.128	.094
4	.126	.107	.216	.141
5	.112	.097	.257	.158
6	.107	.085	.292	.161
7	.105	.076	.321	.173

TABLE 2

Real-World Factor Levels Corresponding to the Coded Levels of the Five Feature Variables

Feature	Levels				
	-2	-1	0	1	2
Number of Edges (NE)	3	4	5	6	7
Area (A)	4000	6000	8000	10000	12000
Radial Length Variance (RV)	M-σ	M-.5σ	M	M+.5σ	M+σ
Edge Length Variance (EV)	M-σ	M-.5σ	M	M+.5σ	M+σ
Major Axis Orientation (MA)	+ 90°	+ 45°	0	-45°	-90°

Generation of Figures

For three of the independent variables any desired level can easily be obtained. The number of edges variable can be specified and the area and major axis orientation can easily be fixed once the figure is generated. For the two variance variables, however, no easy procedure is available to change the value once the figure is generated. Although inefficient, the method used here was to generate a figure, check to see if the variance values obtained were ones that were required. If not, another figure was generated. If the variance values were as required the area and major axis were adjusted to their desired values and the figure was saved on magnetic tape.

The area was adjusted by multiplying all radial lengths by a constant determined by taking the square root of the desired area divided by the actual area. The major axis orientation was fixed by rotating the figure about its centroid. To determine the major axis, the lengths of all lines from a vertex, through the centroid, to the intersecting opposite edge were calculated. The longest of these was determined and the difference between the angle of the associated vertex and the desired angle became the rotation angle.

Two programs were written using this procedure, one for test figures and one for standard figures. The standard figure program required 87 hours to execute and checked over 56,000 figures to obtain the required 810 standard figures. The test figure program was somewhat better requiring only 10 hours to obtain 270 figures.

Generation of Final Stimulus Tape

Having obtained two tapes, one with test figures and one with standard figures, it was necessary to combine these into a single stimulus tape that could be used during actual experimentation. This step involved getting all figures into "virtual memory," randomly selecting a test figure, three standard figures, and a position on the display for the test figure. Three pseudo-random number generators were used to obtain the random numbers required. Offsets were added

to the coordinates of the raw figures so that one figure occupied each of the four display quadrants. The final stimulus tape was generated in such a way that the figures were always displayed in the same order, regardless of which quadrant contained the test figure. This was done to guarantee that any perceptable delay in the generation of the experimental display would not provide the subject with a clue to which position contained the test figure. Each raw test figure and standard figure was used only once. A typical experimental display is presented in Figure 6. The scale in the plot has been adjusted so that the resolution shown in Figure 6 is the same as that on the display the subject observed.

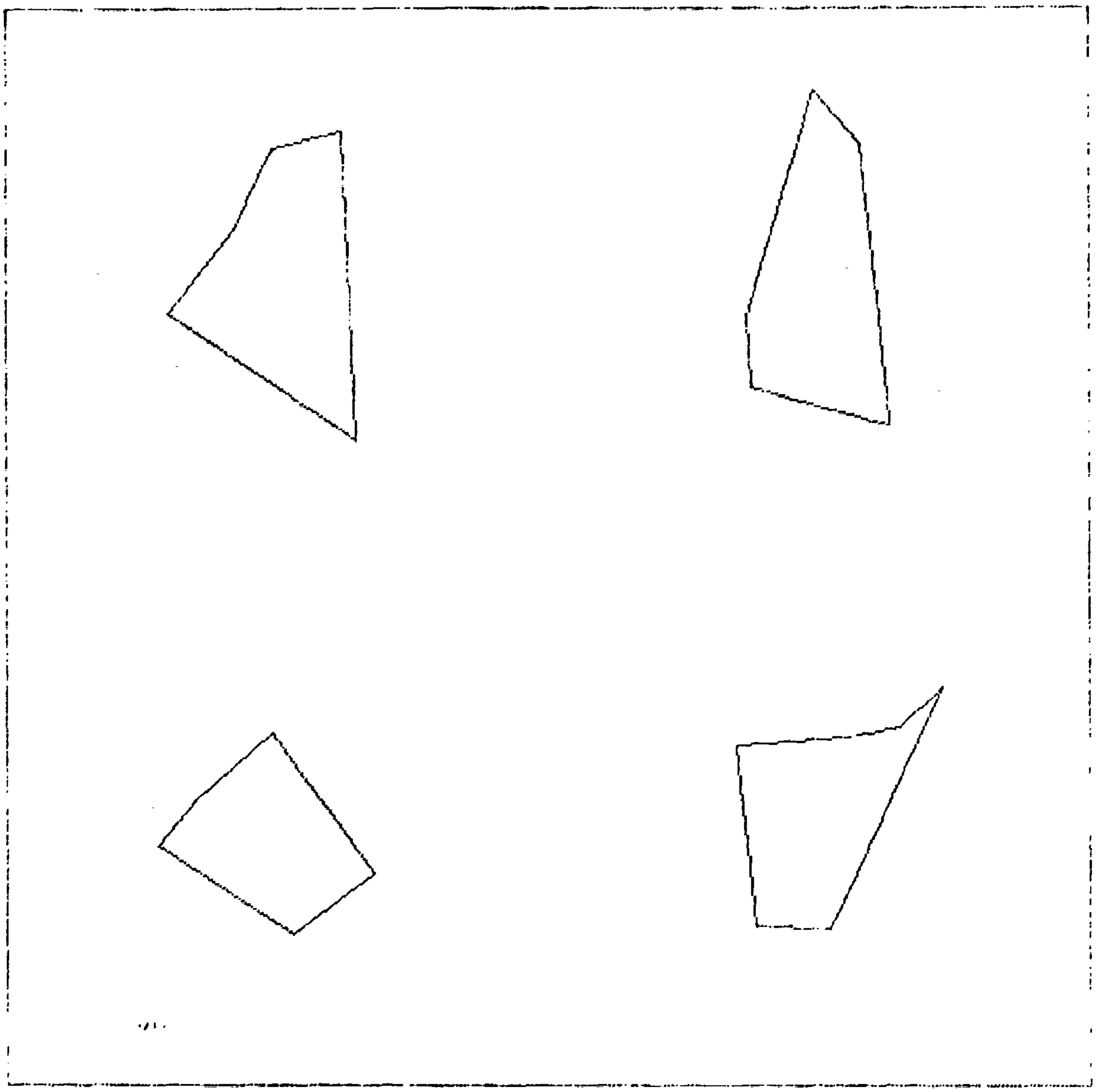


Figure 6. Typical experimental display presented to subjects.

EXPERIMENTAL PROCEDURE

Apparatus

The experiment was computer controlled and used the trial information stored on the stimulus tape. A Digivue plasma panel interfaced to the Raytheon 704 computer of the University of Illinois, Aviation Research Laboratory was used to display the figures. This display is equipped with a touch sensitive input device which was used to obtain the subject's response.

The experimental program was written in Fortran but made extensive use of special Fortran-called Assembly language subroutines to perform special functions. The program was written to allow a completely automatic experiment. Instructions were presented on the plasma panel in text form. When the subject had finished reading, he went on to the experimental trials by touching the panel. The subject could take as long as he desired to make the odd shape discrimination. When he was ready to choose, he merely touched the desired figure. All data were collected on magnetic tape for later analysis.

The experiment was run in two blocks with a short break in between. Half of the subjects saw the first 135 trials first and half saw the second 135 trials first.

Subjects

Ten subjects served in the experiment, 8 male and 2 female. All were members of the staff at the Aviation Research Laboratory and volunteered to be subjects.

ANALYSES AND RESULTS

For each subject and for each of the 27 conditions, the number of times the subject chose the test figure was divided by the total number of trials to yield a probability of choosing the test figure. The probability values were punched on cards along with the appropriate independent variable information. The set of cards was used as input to the RSM analysis program (Clark, Williges, and Carmer, 1971) which produced the following equation:

$$\begin{aligned} PD = .253 &+ .027\,NE - .032\,A + .036\,RV - .013\,EV - .006\,MA \\ &+ .067\,NE^2 + .060\,A^2 - .005\,RV^2 + .036\,EV^2 + .061\,MA^2 \\ &+ .004\,NE\,A - .018\,NE\,RV - .028\,NE\,EV - .024\,NE\,MA + .033\,A\,RV \\ &+ .001\,A\,EV + .037\,A\,MA + .020\,RV\,EV + .012\,RV\,MA - .060\,EV\,MA \end{aligned}$$

where all independent variables are coded values (i.e. -2, -1, 0, 1, 2,) and

PD = Probability of correctly choosing test figure,
NE = Number of Edges,
A = Area,
RV = Radial length Variance,
EV = Edge length Variance, and
MA = Major Axis orientation.

This is the best second-order polynomial multiple regression equation relating the probability of choosing the test figure and the five independent features for the present data. The multiple regression coefficient for this regression equation is .52.

The analysis program also performs an analysis of variance to determine which of the above terms are reliable. That analysis found NE, A, RV, NE^2, A^2, EV^2, MA^2, A RV, A MA, and EV MA to be reliable ($p < .05$). Excluding those terms that were not reliable in the above equation a second analysis provided the following regression equation:

$$PD = .253 + .027\ NE - .032\ A + .036\ RV + .068\ NE^2 + .063\ A^2 + .039\ EV^2 + .064\ MA^2 + .033\ A\ RV + .037\ A\ MA - .060\ EV\ MA$$

The multiple regression coefficient for this second equation is .51. Because the above equation uses coded independent variable values the β weights are directly comparable. The analysis of variance program also partials the residual variance into a subject term and a lack-of-fit term. Both of these terms were highly reliable ($p < .01$). In terms of percent of variance, the regression equation accounts for 23.6 percent, the subject term 7.6 percent and the lack-of-fit term 8.6 percent.

The possibility that subjects were choosing the odd figure based on some unanticipated, but apparent, feature was checked by analyzing common subject responses. For each trial the total number of subjects selecting each of the four possible figures was calculated, and the trials where all subjects made the same response were found. A total of 11 trials were identified. Of these, seven were ones on which all subjects selected the test figure, and four were incorrect choices. An examination of displayed figures for these trials revealed no unusual figures or features.

In interpreting the various terms of the regression equation it is important to keep the characteristics of the experimental task. If a feature being manipulated did indeed influence the subject's selection of odd shape, then the subject responded to the difference in feature level between the standard figures and the test figure. Because the standard figures all had features at the center level, a difference in feature level was represented by the particular level of the test figure feature level regardless of the sign of the coded value. For example, if the difference in

number of edges influenced the subjects selection of odd shape, then both a four- and a six-edged figure would be judged odd with equal probability. The four- and six-edged figures both differ from the standard five-edged figures by one edge even though their coded values are +1 and -1. This sign independence manifests itself in the regression equation by a significant squared term. Four of the five quadratic terms were significant, indicating that the difference between standard figures on these features were apparently being used by subject's in making their choices.

The presence of a linear term indicates that subjects were more likely to select a figure as odd if the feature were in one direction as opposed to the other direction. The addition of a linear term to a quadratic term indicates that not only does the difference in feature level affect choice, but also the direction of the difference. Hence, the significant linear NE term indicates that subjects more often chose a figure with six or seven edges than a figure with three or four edges. The linear by linear terms indicate the interaction of two features.

The intercept term gives the probability of choosing the test figure when all four of the figures have exactly the same features. If these five features adequately describe the figures then the subject must guess in this situation. With four choices the *a priori* probability is .25, and this is the value obtained.

DISCUSSION

The existence of four significant quadratic terms clearly indicates that the number of edges, area, edge length variance, and major axis orientation features predict subject choice of odd shape. Performance was progressively better as the feature became more extreme. For example, the probability of detection equaled .253 when the test figure had its major axis oriented vertically. When it was rotated 45 degrees from the standard figures, the probability increases to .317 but jumped to .509 when the axis was rotated 90 degrees (all other features held constant at 0). This result is consistent with typical psychological studies which demonstrate that discriminability is a function of the magnitude of the physical differences. The probability of detecting a very small difference is quite low; but as the difference becomes larger, the ability to detect the difference improves at an increasing rate. Typically the probability of detection progressively improves until some asymptote is reached after which no further improvement occurs. The value of the asymptote in the most general case is one. When the physical differences reach a particular value, performance is perfect, and additional increases in the difference cannot result in further performance improvement.

The quadratic terms in the regression equation imply that the probability of detection will continue to improve with increasing feature level differences. With a sufficiently large feature difference the regression equation will predict a probability greater than one. This is an obvious physical impossibility and demonstrates two characteristics of the equation obtained. First, prediction beyond the levels of the experimental data is not likely to be valid; and, second, higher-order terms will be required to obtain the expected asymptotic behavior.

Three features had linear effects. In the case of the number of edges feature, the linear effect is positive, indicating that a seven- or six-sided figure is more likely to be chosen as odd. Two possible explanations seem relevant. First, a figure with more edges is generally judged to be more complex (Attneave, 1957; Stenson, 1966) and thus less amenable to simple classification with any of

the other figures. On the other hand, a three-sided figure can easily be labeled as a "triangle" and a couple of the five-sided standard figures classed as "triangle like" resulting in a higher probability of wrong choice. The second explanation is more artifactual. Because it was possible for a figure vertex to occur along the line joining two other verticies, it was possible for a five-sided figure to appear as if it had only four edges. It, of course, was also possible for six-edged figures to appear as five. However, because many more standard figures were presented, the subject was more likely to believe standard figures could have four sides. The possible result is that the number of edges variable was not effective when its value was four. This could then account for the better performance at higher values of NE.

The linear effect of area is negative indicating that figures with smaller areas were more often chosen than those with larger areas. This effect is undoubtably due to the linear scale of area values. The smallest area was half of the area of the standard figures while the largest was only 1.5 times the standard. It is well known that maximum discrimination between areas occurs when they are scaled logrithmically. Hence, the use of the linear scale in the present case made the smaller area much more discriminable.

The third linear term was one for which no quadratic term existed, which indicates that subjects preferred a figure with a larger radial length variance as their choice of odd figure. This can probably be interpreted in terms of the probability of a localized feature such as a notch or teat. When the radial length variance is small the figure tends to be more globular and the chance of getting a distortion quite small. However, when the radial variance is large, the chance of a local feature is much higher. In all probability subjects choose the figure with a distortion as odd.

Three linear by linear interactions terms were also found to be reliable. These were: area by radial length variance, area by major axis orientation, and edge length variance by major axis orientation. The first of these, area and radial

variance, supports the local feature argument above. Performance was best when both the area and the radial variance were large or when both were small. When the area is large, local features can easily be discerned. However, when the figure is small these local features are more difficult to find making radial variance less important. This interaction would also indicate that a large radial variance tends to make a figure appear larger. Indeed, with respect to spatial extention a figure with a large radial variance will be larger, even though the area remains constant. Radial length variance acts to enhance the effectiveness of the area feature.

The interaction of area and major axis orientation appears to be a situation where two cues are better than one. The larger the deviation of the major axis and the area from the standard, the higher the probability of detection. In all probability higher-order interactions of these two variables will also predict performance reliably.

The final interaction term, edge length variance by major axis orientation, has a negative weight indicating that best performance is obtained when the two features are at opposite extremes. With a minimum edge length variance performance is best when the major axis is oriented to the left of vertical. With a large edge variance, figures tilted to the right are preferred. This particular interaction is difficult to understand since there seems to be little reason to believe that a figure lying on its side to the right should be different from a figure on its side to the left.

The regression equation obtained in the present experiment is certainly less than an adequate description of subject performance for a couple of reasons. First, the analysis of variance identified a highly reliable lack-of-fit term, indicating that the second-order equation is not sufficient to describe the data. Second, the reliable subject effect indicates that different regression equations may be necessary for different groups of subjects.

In an attempt to reduce the lack-of-fit, selected higher-order terms were included in the regression analysis. The analysis program allows a specification of the particular terms that will be included in the regression equation. Two

constraints on the specification are immediately apparent. The highest power to which any variable can be raised must be one less than the number of levels of each variable, and the total number of terms must be one less than the number of experimental conditions. For the present study the highest order term would be eighth-order and the maximum number of terms would be 26 (the 27th term is β_0). However, if all 26 terms are specified, no degrees of freedom remain for the lack-of-fit test. The practical limit on the number of terms is, thus, 25.

A serious problem presents itself when trying to fit selected higher-order terms to data collected in accordance with a response surface design. A complete eighth-order polynomial regression equation contains a total of 3124 terms. Problems exist in selecting the specific partial regression terms to fit and in determining when to eliminate a term from the analysis. Similar problems in linear regression have lead to step-wise regression procedures (Ralston and Wilf, 1960) which check all possible combinations of terms. In the present case, however, the number of combinations of 3124 terms taken 25 at a time is staggering thereby making a step-wise procedure impractical.

An investigation of all possible combinations of terms is not necessary. Because the particular RSM design employed consists, in part, of a half-replicate of a 2^5 factorial, a certain amount of intentional confounding exists. The fractional factorial was selected to keep the first- and second-order effects unconfounded, but, higher-order effects are confounded. Because of this confounding, if an improper selection of third- or fourth-order regression term is made the analyses cannot be performed. By examining these confounded effects an identification of those terms that cannot be analyzed together may be obtained. Elimination of those combinations reduces the number of possible regression equations that need be considered.

One of the reasons for checking all possible combinations of terms in a stepwise regression analysis is the possibility that the significance of a particular term will be influenced by the presence or absence of other terms. Due to the characteristics of a RSM design, particular terms will be totally independent of any other terms. This allows these terms to be tested for reliability only once,

and a single decision about whether to keep or drop the term can be made. What remains to be done is an identification of these terms as a function of the characteristics of the experimental design. This identification combined with a development of procedures to select appropriate higher-order terms will greatly enhance the power of RSM designs.

These experimental design investigations were beyond the scope of the present study and consequently no reduction on the lack-of-fit term was accomplished. Pending further RSM investigations, a third-order design will be required on further odd shape studies like the present one.

Another indication that the regression equation obtained could be improved was provided by the reliable subject effect. This effect argues that different subjects attend to different features. This belief is strengthened by the subjects comments upon completion of the experimental session. For example, some subjects said they paid no attention to the orientation of the figure while others indicated that orientation was a strong criteria in their choice. Because each subject performed only once in the experiment, no data were available that would allow a separate regression equation for each subject. Had each subject performed twice, reliability measures on his performance would have been available and individual regression equations would have been possible. With such data it might be possible to separate subjects into several different classes based on the features to which they attended. A separate regression equation for each class would then provide a much better description of subject behavior.

One serious oversight in the present study was a slight confounding of the two variance variables with the number of edges variable. The mean and standard deviation values used when the stimuli were generated were dependent on the number of edges the particular figure possessed. If these two features are to be general characteristics they should be manipulated independently of any other variable. An examination of Table I will reveal that the extent of the confounding is relatively small, however, its existence undoubtably increased the lack-of-fit term. Any further studies should correct this confounding.

CONCLUSIONS

The primary purpose of the present study was to obtain psychological data on human performance that would be suitable for computer simulation. Implicit in this purpose is the necessity for obtaining estimates of the interactive effects of features. The regression equation approach adequately meets this requirement. Given an equation such as the one obtained in this study a possible algorithm for the simulation would be as follows:

1. Calculate the five features for each of the four input figures.
2. For each figure calculate the difference between its feature levels and the average feature levels of the remaining figures.
3. Evaluate the regression equation with these difference values inserted and chose the figure with the largest result as the odd figure.

Because of the lack-of-fit and subject effect problems discussed previously, the particular equation obtained is not as descriptive as would be desired. The present study does, however, indicate the potential usefulness of the regression analysis approach. Chief among the potential benefits is the ability to assess interactions among features. The present study only identified three such terms; however, the inclusion of higher-order terms undoubtably will reveal a great many more.

Additional studies are required to investigate larger ranges of feature values as well as additional features. These studies combined with investigations into the RSM design and analysis problems identified will provide much improved data on human form perception.

REFERENCES

Anastasi, A. The estimation of areas. Journal of General Psychology, 1936, 14, 201-225.

Andrews, M. H., and Brown, D. R. A multidimensional analysis of visual shape discrimination. Paper presented at Midwestern Psychological Association, Chicago, May 1967.

Attneave, F. Dimensions of similarity. American Journal of Psychology, 1950, 63, 516-556.

Attneave, F. Physical determinants of the judged complexity of shapes. Journal of Experimental Psychology, 1957, 53, 221-227.

Attneave, F., and Arnoult, M. D. The quantitative study of shape and pattern perception. Psychological Bulletin, 1956, 53, 452-471.

Bolton, F. E. A contribution to the study of illusions. American Journal of Psychology, 1897-98, 9, 167-182.

Box, G. E. P. and Behnken, D. W. Some new three-level designs for the study of quantitative variables. Technometrics, 1960, 2, 455-475.

Box, G. E. P. and Hunter, J. S. Multifactor experimental designs for exploring response surfaces. Annals of Mathematical Statistics, 1957, 28, 195-241.

Box, G. E. P. and Wilson, K. B. On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, Series B (Methodogical), 1951, 13, 1-45.

Boynton, R. M., Elworth, C. L., Monty, R. A., Onley, J. W., and Klingberg, C. L. Overlap as a predictor of form discrimination under suprathreshold conditions.

Griffis Air Force Base, New York: Rome Air Development Center, Technical Report 61-99, 1961.

Boynton, R. M., Elworth, C. L., Onley, J. W., and Klingberg, C. L. Form discrimination as predicted by overlap and area. Griffis Air Force Base, New York: Rome Air Development Center, Technical Report 60-158, 1960.

Brown, D. R., Hitchcock, L., Jr., and Michels, K. M. Quantitative studies in form perception: An evaluation of the role of selected stimulus parameters in the visual discrimination performance of human subjects. Perceptual Motor Skills, 1962, 14, 519-529.

Brown, D. R. and LoSasso, J. S. Pattern degradation, discrimination difficulty, and quantified stimulus attributes. Psychonomic Science, 1967, 9, 351-352.

Brown, D. R., and Michels, K. M. Quantification procedures, stimulus domains, and discrimination difficulty. Perceptual Motor Skills, 1966, 22, 421-422.

Brown, D. R. and Owen, D. H. The metrics of visual form: Methodological dyspepsia. Psychological Bulletin, 1967, 68, 243-259.

Clark, C. and Williges, R. C. Response surface methodology design variants useful in human performance research. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report ARL-71-21/AFOSR-71-7, October 1971.

Clark, C. and Williges, R. C. Response surface methodology central-composite design modifications for human performance research. Human Factors, 1973, 15, 295-310.

Clark, C., Williges, R. C., and Carmer, S. G. General computer program for response surface methodology analysis. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report ARL-71-8/AFOSR-71-1, May 1971.

Coules, J. and Lekarczyk, M. A. Observer tolerance of form transformation as a function of form complexity. Hanscom Field, Mass.: Electronics Systems Division, Technical Report ESD-TR-63-135, 1963.

Crook, M. N. Facsimile-generated analogues for instrumental displays. In J. W. Wulfeck and J. H. Taylor (Eds.) Form discrimination as related to military problems. Washington, D. C.: National Academy of Science, National Research Council, 1957, 85-98.

Hake, H. W. The study of perception in the light of multivariate methods. In R. B. Cattell (Ed.) Handbook of multivariate experimental psychology Chicago.: Rand McNally, 1966, 502-534.

James, W. The principles of psychology. New York: Holt, 1890.

Koehler, W. Gestalt psychology. New York: Liveright, 1929.

Koffka, K. Principles of gestalt psychology. New York: Harcourt, 1935.

Kunnapos, T., Malhammer, G. and Svenson, C. Multidimensional ratio scaling and multidimensional similarity of simple geometric figures. Scandanavian Journal of Psychology, 1964, 5, 249-256.

Maruyama, K. A problem in form perception: Odd shape detection. Urbana, Ill.: University of Illinois at Urbana-Champaign, Department of Computer Science, Technical Report UIUCDCS-R-71-490, C 00-2118-0027, December 1971.

Mills, R. G. and Williges, R. C. Performance prediction in a single-operator simulated surveillance system. Human Factors, 1973, 15, 327-348.

Monty, R. A. and Boynton, R. M. Stimulus overlap and form similarity under suprathreshold conditions. Perceptual Motor Skills, 1962, 14, 487-498.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Ralston, A. and Wilf, H. S. Mathematical methods for digital computers. New York: Wiley and Sons, 1960, 191-195.

Selfridge, O. G. and Neisser, U. Pattern recognition by machine. Scientific American, 1960, 203, 60-68.

Simon, C. W. The use of central-composite designs in human factors engineering experiments. Culver City, Calif.: Hughes Aircraft Co., Display Systems and Human Factors Department, Technical Report AFOSR-70-6, December 1970.

Stenson, H. H. The physical factor structure of random forms and their judged complexity. Perception and Psychophysics, 1966, 1, 303-310.

Stenson, H. H. Psychological dimensions of simularity among random shapes. Perception and Psychophysics, 1968, 3, 201-214.

Thrumond, J. B. Extensions of an informational deductive analysis of form. Paper presented at the meeting of Southern Society of Philosophy and Psychology, New Orleans, April 1966.

Torgerson, W. S. Theory and methods of scaling. New York: Wiley, 1958.

Warren, J. M. and Pinneau, S. R. Influence of form judgment of apparent area. Perceptual Motor Skills, 1955, 5, 740.

Williges, R. C. and North, R. A. Prediction and cross-validation of video cartographic symbol location performance. Human Factors, 1973, 15, 321-326.

Williges, R. C. and Simon, C. W. Applying response surface methodology to problems of target acquisition. Human Factors, 1971, 13, 511-519.

Zusne, L. Moments of area and of the perimeter of visual form as predictors of discrimination performance. Journal of Experimental Psychology, 1965, 69, 213-220.

Zusne, L. Visual perception of form. New York: Academic Press, 1970.